Making Better Use of Edges via Perceptual Grouping

Yonggang Qi[1], Yi-Zhe Song[2], Tao Xiang[2], Honggang Zhang[1], Timothy Hospedales[2], Yi Li[2], Jun Guo[1]
[1]Beijing University of Posts and Telecommunications. [2]Queen Mary University of London.

Summary: *We propose a perceptual grouping framework that organizes image edges into meaningful structures (See Figure.1) and demonstrate its usefulness on various computer vision tasks. Our grouper formulates edge grouping as a graph partition problem, where a learning to rank method is developed to encode probabilities of candidate edge pairs. In particular, RankSVM is employed for the first time to combine multiple Gestalt principles as cue for edge grouping. Afterwards, an edge grouping based object proposal measure is introduced that yields proposals comparable to state-of-the-art alternatives. We further show how human-like sketches can be generated from edge groupings and consequently used to deliver state-of-the-art sketch-based image retrieval performance. Last but not least, we tackle the problem of free-hand human sketch segmentation by utilizing the proposed grouper to cluster strokes into semantic object parts.*

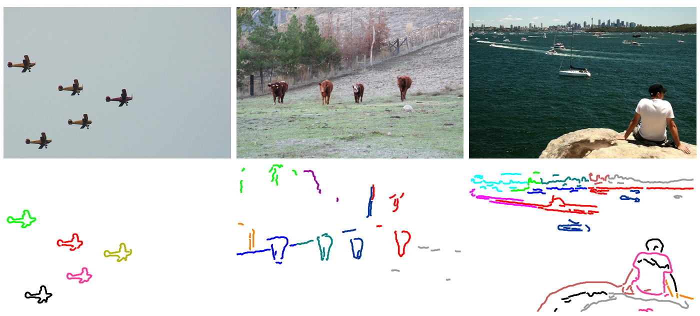

Figure 1: Edge grouping examples (Bottom) on complex scenes (Top).

Grouper: RankSVM is employed to combine continuity and proximity for edge grouping, which aims to learn a ranking function $F(\mathbf{x}) = \boldsymbol{\omega}^T \mathbf{x}$ that outputs a score such that $F(\mathbf{x}(v_i, v_j)) > F(\mathbf{x}(v_i, v_k))$ for any $(v_i, v_j) \succ (v_i, v_k)$. $\mathbf{x}(v_i, v_j)$ is the feature of edge pair (v_i, v_j) and $\boldsymbol{\omega}$ refers to a weight vector adjusting by learning algorithm. Then our grouper formulates edge grouping as a graph partition problem, where the learned RankSVM model is used to encode probabilities of candidate edge pairs as formulated in the following:

$$E(v_L) = \sum_{v_i \in \mathcal{V}} D(v_i, v_L) + \sum_{\{v_i, v_j\} \in N} S(v_i, v_j) \tag{1}$$

$$\begin{aligned} where,\ D(v_i, v_L) &= sigmoid(F(\boldsymbol{x}))^{-1} \\ &= sigmoid(\boldsymbol{\omega}^T \mathbf{x}(v_i, v_L))^{-1} \end{aligned} \tag{2}$$

$$S(v_i, v_j) = d(v_i, v_j)^{-1} \tag{3}$$

Edge grouping for objectness: Inspired by edge-based objectness work[3], a noval criterion to measure objectness is proposed that: (i) the closure area of an edge group should occupy the candidate bounding box as much as possible, since object boundaries often form closed regions, (ii) the structural complexity of edge groups under a candidate box is relatively simple, since ideally each edge group should correspond to exactly one object (or their parts). Therefore, the scoring function is defined in Eq.4. Result is shown in Figure.2 and Tabel.1

$$P_b = \frac{C(G_b)}{n/A_b} = \frac{cvhull(G_b)}{n(b_w \times b_h)^2} \tag{4}$$

Edge grouping for SBIR: Edge grouping is further used to produce human-drawing-like sketches which keep a similar level of details as those from humans, and consequently deliver state-of-the-art sketch-based image retrieval performance as shown in Figure.3 and Table.2.

Edge grouping for sketch segmentation: Edge grouping is capable of clustering sketch strokes into semantic object parts (See Figure.4).

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.

Figure 2: Qualitative examples of our edge grouping based objectness.

Methods	IoU=0.5 Recall(%)	AUC(%)	J_i(%)	IoU=0.7 Recall(%)	AUC(%)	J_i(%)
PercepEdge(Ours)	**95.84**	28.25	79.47	**83.87**	9.83	81.72
PercepEdge-pro	92.25	28.25	78.69	81.97	8.82	80.77
PercepEdge-con	90.53	25.41	78.06	77.70	8.30	80.68
BING	96.39	15.33	65.90	27.34	2.19	78.01
EdgeBoxes50	93.13	21.06	72.61	54.04	5.09	79.42
EdgeBoxes [3]	94.77	28.13	79.68	81.65	10.06	82.32
MCG [1]	93.61	31.44	83.60	77.56	14.08	88.16
Objectness	83.92	15.79	68.81	34.74	2.60	77.48
Sel.Search	91.51	31.04	83.91	77.17	13.96	88.09

Table 1: Comparison of top 10^4 proposals with state-of-the-art on Recall, AUC and Jaccard index at instance level(J_i).

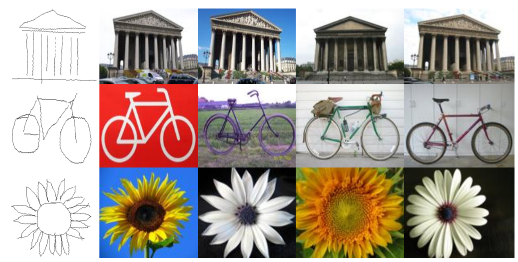

Figure 3: Example query sketch, and their top ranking results.

Methods	Vocabulary size	MAP
PeceptualEdge (Ours)	non-BoW	**0.1837**
PeceptualEdge-proximity	non-BoW	0.1602
GF-HOG [2]	3500	0.1222
HOG	3000	0.1093
SIFT	1000	0.0911
SSIM	500	0.0957
ShapeContext	3500	0.0814
StructureTensor	500	0.0798
PeceptualEdge-continuity	non-BoW	0.0789
StructureTensor	non-BoW	0.0735

Table 2: SBIR results comparison (MAP).

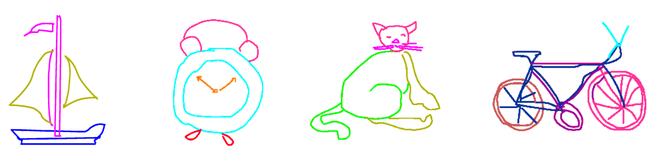

Figure 4: Example sketch segmentation results of four object categories.

[1] Pablo Andrés Arbeláez, Jordi Pont-Tuset, Jonathan T. Barron, Ferran Marqués, and Jitendra Malik. Multiscale combinatorial grouping. In *CVPR*, 2014.

[2] Rui Hu and John P. Collomosse. A performance evaluation of gradient field hog descriptor for sketch based image retrieval. *CVIU 2013*.

[3] C. Lawrence Zitnick and Piotr Dollár. Edge boxes: Locating object proposals from edges. In *ECCV*, 2014.